Distortion of visual localization in three-dimensional virtual space

Sung-en Chien* and Katsumi Watanabe*

*Research Center for Advanced Science and Science and Technology,
The University of Tokyo, Tokyo, Japan.
E-mail: cse@fennel.rcast.u-tokyo.ac.jp Tel: +81-3-5452-5247

Abstract — With applications increasingly using 3D displays, it is important to examine how humans perceive objects located within this medium. The current study examined memory accuracy for the location of a briefly presented visual object when presented using a 3D projector. We observed distortions of depth perception in visual space (i.e., the remembered location of a visual object tended to be underestimated). In addition, we found that this underestimation was not affected by the depth of a fixation point. These results suggest that memorized locations of visual objects tend to be biased toward the observer. We further discuss implications for these findings.

I. INTRODUCTION

Applications using 3D displays have been increasing. People now have more opportunities to interact with visual objects in a virtual 3D space in everyday life. Thus, it is important to examine how humans perceive the location of visual objects in 3D space. The perceived location of a visual object is displaced away from its physical location during steady eye fixations in a 2D plane [1]. A previous study also showed that remembered locations of objects in visual spatial memory are biased toward fixation points in a 2D plane [2]. These studies indicate that visual spatial memory is distorted toward fixation points in 2D space. However, fewer studies have investigated how well people memorize the location of a visual object in 3D space.

Previous studies have found evidence of misperception in 3D space. In a virtual environment, significant underestimation of egocentric distance, the distance from an observer to a visual object, has been shown [3-5]. In these studies, there were no control for observers to fixate upon, and observers were allowed to move their eyes freely while judging distance. Therefore, it is not well known whether the remembered location of a visual object would be biased toward a fixation point or toward the observer in a virtual 3D space.

The current study examined whether remembered depth location is biased toward fixation points (as in a 2D plane) or the observer, regardless of the location of the point of fixation. To this end, visual flashes were presented at different locations in a virtual 3D space. Observers were asked to remember the location of the flash while keeping their eyes fixated on a fixation point during the presentation. The fixation point was presented at different depths to examine whether the remembered location of the flash would be biased toward fixation. If the remembered location of the flash could be biased toward the depth of fixation, the remembered location should be compressed toward the depth of fixation. On the other hand, if the remembered location is biased toward the observer, regardless of fixation, the remembered depth of the flash should be underestimated to the same degree regardless of fixation depth.

II. EXPERIMENT

A. Observers

Four observers participated in the experiment (including one author). All observers had normal or corrected-to-normal visual acuity.

B. Apparatus and stimuli

Visual stimuli were presented via a 3D projector (Sight3D, Solidray) on a large screen (172 cm in width, 130 cm in height) at a viewing distance of approximately 200 cm. Observers viewed visual stimuli through a time-sequential 3D shutter goggle (Nvidia 3D Vision 2), and the refresh rate of the display was 120 Hz (i.e., 60 Hz per one eye). The visual stimuli were presented using a MATLAB operation environment and the Psychtoolbox extensions with OpenGL 3D graphics [6-7]. Calculating binocular disparity generated the images for the left and right eyes. The inter-pupil distance was fixed to 6.5 cm. Visual stimuli appeared in white (3.59 and 55.67 cd/m² for the visual flash) against a black (0.91 cd/m²) background that contained white (4.42 cd/m²) virtual grids on the left and right walls, floor, and ceiling. The spatial interval between each grid was 20 cm. Grids were drawn in the range of 60 cm in front of to 200 cm behind the physical screen.

A white fixation cross was presented at the center of the display. The horizontal distances from the fixation cross to the walls were 50 cm, as were the vertical distances from the center of the fixation cross to the floor and ceiling. The fixation cross could be presented on different depth planes for three experimental conditions. In the 200 depth condition, the fixation cross was presented on the depth plane of the physical screen (200 cm away from observers); in the 240 and 280 conditions, the fixation cross was presented on the depth
plane 40 and 80 cm behind the physical screen (240 and 280 cm away from observers) (Figure 1).

The flash was a white sphere, 1 cm in diameter). However, because no shading was applied, the cue spheres (made in the OpenGL 3D graphics) appeared as 2D disks. In each experimental condition, the flash could be presented at nine possible plane positions. For each possible plane position, the flash might be presented within five possible depth planes ranging from 200 cm to 280 cm away from observers with 20 cm separation. Therefore, there were 45 possible positions for the visual flash (9 plane position × 5 depths; Figures 1, 2).

C. Procedure

Observers began each trial by pressing the space key. The fixation was presented alone for 1000 ms, and observers were instructed to focus on the fixation until the flash disappeared. The flash as a target was then presented for 100 ms and disappeared. After 1,000 ms, a white sphere identical to the target (as the probe) appeared at an unpredictable location in the virtual space. Observers were instructed to move the probe to the remembered location of the flash via a 3D mouse, allowing observers to move the sphere freely in the virtual 3D space. The probe could be moved in depths from 120 cm to 400 cm away from observers and every possible plane location within the virtual grid. Observers were allowed to move their eyes during the response phase. Observers pressed the space key again to finish a test trial once they had moved the probe to the remembered location of the flash (Figure 3). The final location of the probe was recorded. In each experimental condition with the fixation presented at different depth planes, there were 135 test trials (9 plane positions × 5 depths × 3 repetitions). In addition to the experimental conditions, there was a control condition to account for motor response errors. In the control condition, the target would not vanish. Observers were instructed to move the probe to the same position as the target. The number of test trials was identical to the experimental conditions. Each observer performed the four conditions in a random sequence.

D. Results

We calculated the average “deviation in depth” to examine whether or not the remembered depths were biased. For every possible flash position, the true distance of the flash was subtracted from the average recorded distance in the control condition as a baseline. We then subtracted the baseline and true distance of the flash from the average remembered distance in each experimental condition to obtain the
deviation in depth. Therefore, positive deviation values indicated that the flash depth was underestimated and vice versa.

The results are shown in Figure 4. A repeated measures two-way analysis of variance (ANOVA) revealed that the main effect of flash depth was significant ($F(4,12) = 3.873, p < .05$), implying that when the flash was presented further away from the observer, the underestimation of distance was larger. However, the main effect of depth position for the fixation cross was not significant ($F(2,6) = 1.420, p = .313$). There was also no interaction between fixation position and flash position ($F(8,24) = .697, p = .691$).

The results showed that the remembered positions’ visual flash depth was not compressed toward a fixation point but biased toward observers regardless of fixation depth. While previous studies suggest that visual spatial memory of plane positions is biased toward a fixation in 2D space [1-2], our results imply that visual spatial memory for depth positions is not biased in the same manner but toward observers.

III. DISCUSSION

The present study examined whether remembered depth location is biased toward a fixation based on plane position or the observer, regardless of fixation position, in a 3D space. The results implied that the remembered location is underestimated toward the observer and not affected by fixation depth.

Previous studies of spatial memory for visual objects in a 2D plane often observed a bias toward fixation. One possible interpretation is that the fixation point can serve as a salient landmark. This is supported by studies showing that the remembered locations of both moving and stationary objects are biased toward a salient visual object presented at a location other than a fixation point [9-10]. Additionally, the compression of visual spatial memory toward a fixation point is weaker when a salient visual reference is presented in the display [1]. These studies suggest that visual memory is constructed from information related to visual references in space.

Spatial memory for depth position might have different characteristics compared to spatial memory for a 2D plane position. Depth perception is also constructed by depth cues from the environment. However, for humans, the basic function of depth perception is to estimate the distance between visual objects and the self. Therefore, humans can deal with dynamic changes in the environment, such as the approach of potential dangers. Along this line, the most important reference for constructing spatial memory of depth might be our own selves or bodies. This might be the reason that remembered depth positions were consistently biased toward the observer in the present study.

The current study examined how well people could memorize depth position in a virtual 3D space. However, several questions remain. Given that our environment changes dynamically, future studies should focus on examining the effects of moving visual stimuli or dynamic backgrounds in constructing visual spatial memory.